Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER

2. GOVT ACCESSION NO. AD-A109 046

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)

Random Indeoendence Systems

5. TYPE OF REPORT & PERIOD COVERED

Technical Report

6. PERFORMING ORG. REPORT NUMBER

TR-32-81

7. AUTHOR(s)

John H. Reif
Paul G. Spirakis

8. CONTRACT OR GRANT NUMBER(s)

N00014-80-C-0674

NSF

9. PERFORMING ORGANIZATION NAME AND ADDRESS

Harvard University
Cambridge, MA 02138

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS

Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

12. REPORT DATE

November, 1981

13. NUMBER OF PAGES

21

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

same as above

LEVEL II

15. SECURITY CLASS. (of this report)

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

unlimited

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

independence system, matroid, second moment method, combinatorial optimization, random graph.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report ~~We~~ introduces a new random structure generalizing matroids. These *random independence systems* allow us to develop general techniques for solving hard combinatorial optimization problems with random inputs.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

Unclassified 405331

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Center for Research in Computing Technology

RANDOM INDEPENDENCE SYSTEMS

by

John H. Reif
Paul G. Spirakis

TR-32-81

RANDOM INDEPENDENCE SYSTEMS*

John H. Reif and Paul G. Spirakis

Aiken Computation Laboratory
Harvard University
Cambridge, Mass.

November, 1981

*A previous draft of this paper, titled "Random Matroids," was presented at the 1980 ACM Symposium on Theory of Computing which was held in Los Angeles, California.

*This work was supported in part by the National Science Foundation Grant NSF-MCS79-21024 and the Office of Naval Research Contract N00014-80-C-0674.

ABSTRACT

We introduce a new random structure generalizing matroids. These *random independence systems* allow us to develop general techniques for solving hard combinatorial optimization problems with random inputs.

1. Introduction

In a classic paper "On the Abstract Properties of Linear Dependence" of 1935, Whitney provided a set of axioms for a structure commonly called a *matroid*. Matroid theory (see [Tutte, 1971], [Lawler, 1976]) has applications to a wide class of combinatorial optimization problems: where we wish to construct a maximal object (a maximum independent set) satisfying a monotone property. Intersections of matroids, are called *independence systems* (see [Korte, Hausmann, 1978]) and have also wide practical applications. The problem of constructing a maximum independent set in an independence system is NP-complete for independence systems which are intersections of three or more matroids.

We introduce in this work (Section 2) the *random independence system* (RIS), which is applicable to a more general class with combinatorial optimization problems with *random inputs*. We define some natural notions, such as "maximal with a given probability."

Section 3 sketches a general *nonconstructive proof technique* for determining the existence (with probability 1) of an independent set of given cardinality in instances of an RIS. This encompasses various nonconstructive proofs of graph properties in [Erdös and Spencer, 1947] and uses the second moment method.* In Section 4 we define a *weighted*

*In contrast, a companion paper [Reif, Spirakis, 1981], discussed the construction of random independent sets by use of an extension-rotation algorithm.

RIS. We also provide a nonconstructive proof technique for determining the existence of an independent set of given weight in a weighted RIS and a result on the relation of the existence of maximum independent sets in a RIS to the values of the weights of maximum independent sets in a weighted RIS. Section 6 discussed intersections of RIS and the relationship of RIS to matroids.

2. Definitions of Random Independence Systems and Their Structure

2.1 Definitions of Random Independence Systems

Let E be a set and let $\mathcal{J}$ be a family of subsets of E. For each element $e \in E$, let the element's probability be a real number p_e defined on the interval $[0,1]$. The triple $M = (E, \mathcal{J}, \{p_e\})$ is a *random independence system* (RIS). If for some fixed p, $p_e = p$ for all elements $e \in E$, then M is *uniform* and denoted $(E, \mathcal{J}, p)$. We will frequently write $(E, \mathcal{J}, 1)$ as $(E, \mathcal{J})$. M is a *proper RIS* if

(A1) $\emptyset \in \mathcal{J}$

(A2) $A \in \mathcal{J} \wedge A' \subseteq A \Rightarrow A' \in \mathcal{J}$.

Intuitively, $\mathcal{J}$ may be considered a property on subsets of E which is *trivially satisfied* (by axiom A1) and *monotone decreasing* (by axiom A2). A pair $(E, \mathcal{J})$ satisfying A1 and A2 is called a (deterministic) independence system (see [Korte, Hausmann, 78]). Let A3 be the axiom: for any sets $A, A' \in \mathcal{J}$ of cardinality h, $h+1$ respectively, $\exists e \in A' - A$ such that $A \cup \{e\} \in \mathcal{J}$. If $(E, \mathcal{J})$ satisfies the axioms A1, A2 and the additional axiom A3, then it is commonly called a *matroid* (see [Whitney, 1935]). (Note that an alternative axiom to A3 is "for any subset $S \subseteq E$,

all maximal independent subsets of S have the same cardinality").

2.2 Instances of Random Independence Systems

An *instance* of a random independence system $M=(E,\mathcal{J},\{p_e\})$ is a pair $M_0=(E_0,\mathcal{J}_0)$ where

(i) $E_0 \subseteq E$ is derived by independently choosing each $e \in E$ with probability p_e

(ii) $\mathcal{J}_0=\{I \in \mathcal{J}/I \subseteq E_0\}$.

Note that the probability of M_0 is

$$\left(\prod_{e \in E_0} p_e\right)\left(\prod_{e \in E-E_0} (1-p_e)\right) .$$

Clearly, any instance M_0 of a proper RIS satisfies axioms A1 and A2 and thus it is a (deterministic) independence system.

A set $A \subseteq E_0$ is independent in M_0 if $A \in \mathcal{J}_0$ and *dependent* otherwise. An independent set $I \in \mathcal{J}_0$ is *maximum* in M_0. if there does not exist any $I' \in \mathcal{J}_0$ such that $|I'|>|I|$. Let the *rank* of M_0 be the cardinality of maximum independent set. $I \in \mathcal{J}_0$ is *maximal* in M_0 if there does not exist any $I' \in \mathcal{J}_0$ such that $I' \supset I$. A *minimal dependent set* of M_0 (a *circuit*) has no proper subset which is dependent in M_0. For any $A \subseteq E_0$ let the *rank of* A *in* M_0 be the maximum cardinality of any independent subset of A.

2.3 Examples of RIS

As an example of a RIS, let Q be a property on graphs and let $G_{n,p}$ be a random undirected graph. $G_{n,p}$ has instances which are graphs with vertices $V=\{1,2,\ldots,n\}$ and each edge chosen independently with

probability p from unordered pairs of distinct vertices in V. Let $M = (E, \mathcal{J}, p)$ be the uniform RIS with $E = \{\{u,v\}/\text{distinct } u,v \in V\}$ and $\mathcal{J} = \{E' \subseteq E/Q(V,E') \text{ holds}\}$. Then any instance $M_0 = (E_0, \mathcal{J}_0)$ of M corresponds to an instance (V,E_0) of the random graph $G_{n,p}$ and $\mathcal{J}_0$ contains precisely those edge sets $I \subseteq E_0$ such that the property Q holds for the subgraph (V,I). The graph property Q is *trivially satisfied* (Q holds for the graph with no edges) and *decreasing monotone* (i.e., $Q(G) \Rightarrow Q(G')$ for all subgraphs G' of G) iff M is a proper RIS.

(a) Given a graph $G = (V,E)$ a *simple path* is a path of edges in E containing no cycles, and it is a *Hamiltonian path* if it contains every vertex of V. The property of a "simple path" in a random graph does not yield a proper RIS, since a simple path must be connected (violating axiom A2). However, we can define a proper RIS such that any independent set of cardinality $|V| - 1$ is a Hamiltonian path. We give both formulations here:

Formulation as a non-proper RIS: Let $M = (E, \mathcal{J}, p)$ be the RIS where $\mathcal{J}$ is the set of all simple paths in the complete graph (V,E). Fix an instance $M_0 = (E_0, \mathcal{J}_0)$ of M. Then (V,E_0) has the same probability in random graph $G_{n,p}$ as in M and $\mathcal{J}_0$ is the set of all simple paths in (V,E_0).

Formulation as a proper RIS: Let $M = (E, \mathcal{J}, p)$ be the RIS with E as above and $\mathcal{J} = \{I \subseteq E/(V,I) \text{ consists of a set of disjoint sample paths}\}$.

Clearly M satisfies axioms A1, A2. Fix an instance $M_0 = (E_0, \mathcal{J}_0)$ of M. Then (V,E_0) has the same probability in $G_{n,p}$ as in M and

$\mathcal{J}_0$ has as elements all different set of disjoint simple paths in E_0.

In both formulations, if M_0 has an independent set $I \in \mathcal{J}_0$ such that $|I| = |V| - 1$ then (V,I) is a Hamiltonian line in (V,E).

(b) An edge *matching* of a graph is a set of vertex disjoint edges and is *perfect* if every vertex appears in some edge of the matching. To formulate the "perfect matching" problem as an RIS, we assume a complete graph $G = (V,E)$ with $2n$ vertices. Let $M = (E,\mathcal{J},p)$ where $\mathcal{J} = \{I \subseteq E / I$ is a matching$\}$. Let $M_0 = (E_0, \mathcal{J}_0)$ be an instance of M. Then M_0 has a perfect matching if there is an $I \in \mathcal{J}_0$ such that $|I| = n$. The property of "matching" in a random graph $G_{2n,p}$ yields a proper RIS, since if I is a matching, then every $I' \subseteq I$ is a matching. A subgraph $G' = (V',E')$ of a graph $G = (V,E)$ is called a *clique* if $E' = \{\{u,v\}/u,v$ distinct vertices of $V'\}$. The *clique* property in random graphs $G_{n,p}$ gives a *proper* RIS $M = (V,\mathcal{J},p)$ where $\mathcal{J} = \{I \subseteq V/I$ is the vertex set of a clique in $G = (V,E)\}$.

(c) An r-*coloration* of a graph G is a map $h: G \to \{1,2,\ldots,r\}$ such that for all edges $e = \{u,v\}$ of G, $h(u) \neq h(v)$. The minimal r for which such an r-coloration exists is called the chormatic number of G and denoted by $X(G)$.

The r-*coloration property* in random graphs $G_{n,p}$ gives a *proper RIS* where

$\mathcal{J} = \{I \subseteq E/$the subgraph of G induced by I is r-colorable$\}$.

(d) A *complete* k-*partite subgraph* of the graph $G = (V,E)$ is defined by a collection $\{V_1,\ldots,V_k\}$ of pairwise disjoint subsets of V such that $\{u,v\} \in E$ for $u \in V_g$, $v \in V_h$ iff $g \neq h$. The complete k-partite subgraphs of instances of $G_{n,p}$ correspond to the independent

sets of instances of the following RIS: $M=(E,\mathcal{J},p)$ where $I\in\mathcal{J}$ if there exists a partition $p(I)=\{V_1,\dots,V_k\}$ of I (that is $\bigcup_{h=1}^{k} V_h=I$ and $V_h\cap V_g=\emptyset$ for $1\le g<h\le k$) that defines a complete k-partite graph on I. This RIS is *proper* too.

(e) Let $V_1=\{1,\dots,n\}$, $V_2=\{n+1,\dots,2n\}$ be disjoint vertex sets of equal cardinality and let $E=\{\{u,v\}/u\in V_1, v\in V_2\}$. A *bipartite graph* $B=(V_1\cup V_2,E_0)$ has vertex set $V_1\cup V_2$ and edge set $E_0\subseteq E$. B is *complete* if $E_0=E$. A *random bipartite graph* $B_{n,p}$ has instances which are bipartite graphs $(V_1\cup V_2,E_0)$ where each edge of E_0 is chosen from E with probability p.

An *(edge) matching* of bipartite graph $(V_1\cup V_2,E_0)$ is a set of vertex disjoint edges $I\subseteq E_0$ and is *perfect* if every vertex of $V_1\cup V_2$ appears in some edge of I. The *bipartite perfect matching problem* is formulated as a *proper* RIS by assuming a complete bipartite graph $B=(V_1\cup V_2,E)$. Let $M=(E,\mathcal{J},p)$ where $\mathcal{J}=\{I\subseteq E/I$ is a (bipartite) matching$\}$. Let M_0 be an instance of M. M_0 has a perfect matching if there is an $I\in\mathcal{J}_0$ such that $|I|=n$.

(f) Let S be a finite set, $|S|=n$, and let $E=\{S_1,\dots,S_m\}$ be a family of subsets of S. A subfamily $I\subseteq E$ is a *packing* in S if the sets in I are pairwise disjoint. Let $M=(E,\mathcal{J},p)$ be the *proper uniform* RIS whose $\mathcal{J}=\{I/I$ is a packing$\}$. An instance of M correspond to an instance of a random hypergraph having vertex set S and obtained by selecting each of $S_1,\dots,S_m$ with equal probability p (independently). Maximal packings in the instances, covering all "vertices" of S correspond to independent sets $I\in\mathcal{J}_0$ such that the union of the elements of I gives S.

(g) Suppose that there are n courses possible to be taken by a student, and they have to be done one at a time, starting at time 0. Each course $i \in \{1,\ldots,n\}$ has a fixed uninterrupted duration time ti and deadline di. Let a *random course assignment* $A_{n,p}$ be the random variable whose instances are subsets E_0 of $\{1,\ldots,n\}$ created by choosing each i with equal probability p, independently. Let $I \subseteq E_0$ be a *proper* assignment if all courses in I can be completed by their deadlines. The property of *proper assignment* gives a *proper uniform* RIS $(E,\mathcal{J},p)$ with $\mathcal{J} = \{I/I \subseteq \{1,\ldots,n\}$ and I is proper assignment$\}$.

(h) Consider a set $V = \{X_1,\bar{X}_1,\ldots,X_N,\bar{X}_N\}$ of literals. Let us choose each of the elements of $E = V \times V \times V$ independently with probability p. The subset obtained can be viewed as a Boolean expression in 3-conjunctive normal form. The random variable whose instances are described by the above experiment, is called a *random Boolean Expression*, $BOOL_{N,p}$. A subset I of an instance of a random boolean expression is called *satisfiable* if there is an assignment of exactly one of the values {true, false} to each of the literals appearing in I (such that if both X_j, $\bar{X}_j$ appear then $\text{value}(X_j) = \neg\text{value}(\bar{X}_j)$) and such that the evaluation of I gives *true* as an answer. The *proper uniform* RIS $M = (E,\mathcal{J},p)$ where $\mathcal{J} = \{I \subseteq E/I$ satisfiable$\}$ has instances corresponding (having the same probability) with the instances of $BOOL_{N,p}$. An instance of $BOOL_{N,p}$ is satisfiable iff the corresponding $(E_0,\mathcal{J}_0)$ of M has an independent set of cardinality $|E_0|$.

2.4 Maximality in Random Independence Systems

The definitions of maximum, maximal and minimal are all standard for monotone properties of *deterministic* combinatorial structures. We extend these notions to independence in RIS which is a *random* property.

Let $M = (E, \mathcal{J}, \{p_e\})$ be an RIS and let $A \in \mathcal{J}$. Let A be *maximum with probability* m in M if

$$m = \text{Prob}\{A \text{ is maximum in the same instance}/ A \text{ appears in an instance}\}.$$

(All probabilities are defined over the possible instances of M.)

Let A be *maximal with probability* m *in* M if

$$m = \text{Prob}\{A \text{ maximal in the same instance}/ A \text{ appears in an instance}\}.$$

Similarly, let a set $A \in 2^E - \mathcal{J}$ be *minimal* (a *circuit*) *with probability* m *in* M if

$$m = \text{Prob}\{A \text{ is a minimal dependent set of the same instance}/ A \text{ appears in an instance}\}.$$

Let *rank*(M) be the random variable giving the rank of instances of M. For all $m \in [0,1]$, let $\delta_M(m)$ be the minimal $m' \in [0,1]$ such that $\forall A \in \mathcal{J}$: A maximal with probability m in M implies A is maximum with probability $\leq m'$ in M. (It is obvious that $m \geq m'$ and that $\delta_M(m)$ is increasing with p.)

The function $\delta_M(m)$ gives us a measure with which simple greedy-like algorithms succeed in constructing maximum sets. A similar function

may be defined for the measure of success of rotation-extension algorithms. (See [Reif, Spirakis, 1981]). Note that for matroids $\delta_M(m) = m$.

3. A General Nonconstructive Existence Theorem

Let $M = (E, \mathcal{J}, p(\ell))$ be a uniform RIS where $\ell = |E|$. Let $\mathcal{J}_h = \{I \in \mathcal{J} / |I| = h\}$ for $h \geq 1$. Let *the interdependence ratio for* M be

$$IR_h = \text{the mean of } \frac{\text{Prob}\{I \text{ independent}/I' \text{ independent}\}}{\text{Prob}\{I \text{ independent}\}}$$

for $I, I' \in \mathcal{J}_h$.

For a fixed $h > 0$, we are interested in a minimum $p(\ell)$ (the critical p) such that as $\ell \to \infty$

$$\text{Prob}\{\text{rank}(M) \geq h\} \to 1$$

or equivalently

$$\text{Prob}\{\exists \text{ independent set of size } h \text{ in any instance of } M\} \to 1 \quad \text{as } \ell \to \infty .$$

The following is a generalization of a nonconstructive proof technique due to Erdös and Renyi.

THEOREM 3.1. *If for* $\ell \to \infty$, $IR_h = 1 + o(1)$ *then the critical* p *is lower bounded by*

$$|\mathcal{J}_h|^{-1/h} \quad \textit{for} \quad |\mathcal{J}_h| > 0$$

Proof. Let I range over the members of $\mathscr{I}_h$ and let X_I be the random variable being 1 if I is independent in an instance M_0 and 0 else, for each instance M_0 of M. Let $Y = \Sigma X_I$, I ranges over $\mathscr{I}_h$.

It is clear that $Y > 0 \Rightarrow \text{rank}(M) \geq h$.

From the Chebyshev inequality

$$\text{Prob}(Y=0) \leq \frac{\text{Var}(Y)}{\text{mean}^2(Y)} .$$

The

$$\text{mean}(Y) = \sum_{I \in \mathscr{I}_h} \text{mean}(X_I) = p^h |\mathscr{I}_h| .$$

The variance of Y is

$$\begin{aligned} \text{Var}(Y) &= \text{mean}(Y^2) - \text{mean}^2(Y) \\ &= \sum_{\substack{I,J \in \mathscr{I}_h \\ I \neq J}} \text{mean}(X_1 X_J) - \text{mean}^2(Y) . \end{aligned}$$

But

$$\begin{aligned} \text{mean}(X_I X_J) &= \text{Prob}\{X_I = X_J = 1\} \\ &= \text{Prob}\{X_J = 1/X_I = 1\} \cdot \text{Prob}\{X_I = 1\} \\ &= IR_h \cdot \text{mean}^2(X_I) \\ &= IR_h \cdot (p^h)^2 . \end{aligned}$$

Also,

$$\text{mean}^2(Y) = (p^h)^2 \cdot |\mathscr{I}_h|^2 .$$

Hence

$$\text{Prob}(Y=0) \leq \frac{\text{mean}^2(Y)(IR_h - 1)}{\text{mean}^2(Y)}$$

or

$$\text{Prob}(Y=0) = o(1) \quad \text{as} \quad \ell \to \infty .$$

Since we also want mean$(Y) \geq 1$ we get $p \geq |\mathcal{J}_h|^{-1/h}$. □

In practice, the bound $p \geq |\mathcal{J}_h|^{-1/h}$ may not be sufficient to guarantee $IR_h = 1 + o(1)$ as $\ell \to \infty$. To compute IR_h, we introduce a new random variable $u = |I \cap J|$ for randomly chosen $I, J \in \mathcal{J}_h$. Then

$$IR_h = \text{mean}(p^{h-u}/p^h) = \text{mean}(p^{-u})$$

$$= \sum_{k=0}^{h} p^{-k} \cdot \text{Prob}\{u=k\} .$$

Thus we must choose p to satisfy also

$$\sum_{k=0}^{h} p^{-k} \cdot \text{Prob}\{u=k\} = 1 + o(1) \quad \text{as} \quad \ell \to \infty .$$

In fact, we want the probability of large intersections to be small. This is formalized in the following theorem:

THEOREM 3.2. *If*

$$\text{Prob}\{u=k\} \leq p^{2k}(1-p)(1+o(1))$$

then

$$IR_h = 1 + o(1)$$

and thus the critical p *found in the previous theorem suffices.*

Proof. Since u is an integer then $p^{-u} \geq 1$ and hence

$$IR_h \geq 1 \;. \tag{a}$$

Also

$$IR_h = \sum_{k=0}^{h} p^{-k} \, \mathrm{Prob}\{u = k\}$$

$$\leq \sum_{k=0}^{h} p^{-k} p^{2k} (1-p)(1+o(1))$$

$$\leq (1-p)(1+o(1)) \cdot \frac{1-p^{h+1}}{1-p} < 1 + o(1) \tag{b}$$

By (a) and (b), $IR_h = 1 + o(1)$. □

Let us consider RIS $M = (E, \mathcal{J}, p(\ell))$ for various properties of random graphs, of the model $G_{n,p}$ with n vertices V and $\ell = \frac{n(n-1)}{2}$ possible edges $E = \{\{u,v\} | u,v \in V\}$, each chosen with probability p. For *clique* of c vertices and $h = \frac{c(c-1)}{2}$ edges, $|\mathcal{J}_h| = \binom{n}{c}$ and the critical p is 1/2 for $h = 2 \log n$, derived directly from Theorem 3.1. For *perfect matchings* of h edges, $|\mathcal{J}_h| = \binom{n}{h}\binom{n-h}{h} h!$ and the critical p is $\theta(\frac{\log n}{n})$ for $h = n/2$ (n even). It is again derived directly from Theorem 3.1.

For a *Hamiltonian path* of h edges $|\mathcal{J}_h| = h!\binom{n}{h+1}$ and the critical p is $\theta(\frac{\log n}{n})$ for $n = n-1$. It was derived by Posa [1976] by a constructive technique (generalized in [Reif, Spirakis, 1981]). Theorem 3.1 does not seem applicable in this case. On the other hand, there is no known efficient (polynomial time) algorithm for constructing cliques of size $2 \log n$ with probability $\to 1$ when the edge density is the critical $p = 1/2$.

For *maximal packings* in case of $|S_1| = |S_2| = \cdots = |S_m| = 3$ we get

$$|\mathcal{P}_h| = \binom{n}{h}\binom{n-h}{h}\binom{n-2h}{h}(h!)^3$$

and the critical p is $n^{-3/2}\cdot\alpha(n)$ for $h = n/3$ and $\alpha(n) \to \infty$ as $n \to \infty$. (The same p gives perfect matchings in 3-regular hypergraphs.)

4. Weighted RIS

We now extend our definition of RIS so that the elements are independently, randomly weighted over given probability distributions. We wish upper and lower bounds on the weight of the maximum independent set. Lueker [1978] considers this problem for graphs with a normal distribution of edge weights and we show his results extend to weighted RIS with arbitrary uniform distributions.

A *weighted RIS* M is a triple $(E, \mathcal{J}, \{w_e\})$ where E is a set of elements, $\mathcal{J} \subseteq 2^E$ and for each $e \in E$, W_e is some independent random variable.

An *instance* of M is $M_0 = (E, \mathcal{J}, \{W_e\})$ where the W_e are instances of the w_e for each $e \in E$. M is uniform if the W_e have the same distribution.

Let $\mathcal{J}_{max}$ be the set of all maximum (in cardinality) elements of $\mathcal{J}$. Let h_0 = size of maximum elements of $\mathcal{J}$. For all $I \in \mathcal{J}$, let $W(I) = \Sigma_{e \in I} W(e)$. Let $W_{max}(M)$ be the random variable

$$\max\{W(I)/I \in \mathcal{J}_{max}\} \quad .$$

We immediately get

PROPOSITION 4.1. mean($W_{max}(M)$) $\leq$ mean *of* $W(I)$ *over all* $I \in \mathscr{I}_{max}$ *and instances of* M *such that* $W(I) = W_{max}(M)$. *For example, if the* $\{W_e\}$ *are all normal with mean* μ *and variance* σ^2, *then*

$$\text{mean}(W_{max}(M)) \leq h_0\mu + \sigma\sqrt{2h_0 \log|\mathscr{I}_{max}|} \qquad as \quad |E| \to \infty . \qquad \square$$

Let $M = (E, \mathscr{I}, W)$ be a uniform weighted RIS and let $\ell = |E|$. Let F be the probability distribution of W and choose some $p \in (0,1)$. For any instance $M_0 = (E_0, \mathscr{I}_0, W)$ of M, let $M_0' = (E_0', \mathscr{I}_0')$ be derived from M_0 by deleting each element $e \in E$ with $W_e < F^{-1}(1-p)$ and let $\mathscr{I}_0' = \{I \in \mathscr{I}_0 / I \subseteq E_0'\}$. We claim that instances of $M' = (E, \mathscr{I}, p)$ have the same probability as the corresponding M_0' instances. To see that, note that an element e is chosen with probability

$$p' = \text{Prob}\{W_e > F^{-1}(1-p)\} = 1 - F(F^{-1}(1-p)) = p .$$

Thus,

PROPOSITION 4.2.

$$\text{mean}(W_{max}(M)) \geq |\mathscr{I}_{max}| \cdot F^{-1}(1-p)$$

if

$$\text{mean}(W_{max}(M)/\text{rank}(M') < h_0) = o(|\mathscr{I}_{max}| \cdot F^{-1}(1-p))$$

as $\ell \to \infty$. $\square$

Note that if the restriction of Proposition 4.2 is satisfied, we have an algorithm which with high likelihood (as $\ell \to \infty$) constructs an independent set with weight $\geq |\mathscr{I}_{max}| \cdot F^{-1}(1-p)$ in an instance of M. This idea has been used by Walkup [1977] for discrete distributions of

W and by Lueker [1978] for W with normal distributions.

For example, assume W is normal with mean μ and variance σ^2 and let $q = \text{Prob}\{\text{rank}(M') = h_0\}$. Then if

$$q\sqrt{-h_0 \log q} = o(h_0\sqrt{-\log q})$$

then

$$\text{mean}(W_{max}(M)) \geq h_0 \cdot \mu + h_0 \cdot \sigma\sqrt{-2 \log p} \quad .$$

5. Nonconstructive Existence Theorem for a Weighted RIS

Next we describe a nonconstructive existence proof technique for weighted RIS. Let $M = (E, \mathcal{J}, W)$ be a weighted RIS where W is a mapping from E to the positive reals and let $\ell = |E|$. Let $\mathcal{J}_{max}$ be the sets of $\mathcal{J}$ of maximum cardinality and let for every $I \in \mathcal{J}_{max}$, X_I^k be the random variable

$$\begin{aligned} X_I^k &= 1 \quad \text{if} \quad W(I) \geq k \\ &= 0 \quad \text{else} \quad . \end{aligned}$$

Let

$$Y_k = \sum_{I \in \mathcal{J}_{max}} X_I^k$$

and let the *weight interdependence ratio* be

$$\text{WIR}_k = \text{mean}\left(\frac{\text{Prob}\{W(I) \geq k / W(J) \geq k\}}{\text{Prob}\{W(I) \geq k\}}\right)$$

for $I, J \in \mathcal{J}_{max}$.

Again, we can prove as in Section 2 that

$$\text{mean}(Y) = |\mathcal{J}_{max}| \cdot \text{Prob}\{I \in \mathcal{J}_{max} \text{ has } W(I) \geq k\}$$

and

$$\text{Var}(Y) = \text{mean}^2(Y)(\text{WIR}_k - 1) .$$

By the Chebyshev Inequality,

$$\text{Prob}\{Y = 0\} < \frac{\text{Var}(Y)}{\text{mean}^2(Y)} = \text{WIR}_k - 1 .$$

So, if

$$|\mathcal{J}_{max}| \cdot \text{Prob}\{I \in \mathcal{J}_{max} \text{ has } W(I) \geq k\} \geq 1$$

and

$$\text{WIR}_k \to 1 + o(1) \quad \text{for} \quad \ell \to \infty$$

then

$$\text{Prob}\{Y = 0\} \to 0 \quad \text{as} \quad \ell \to \infty$$

or

$$\text{Prob}\{\exists I \in \mathcal{J}_{max} \text{ with weight } \geq k\} \to 1 \quad \text{as} \quad \ell \to \infty .$$

By the Central Limit Theorem, we get:

THEOREM 5.1. *Suppose that* M *is uniform, so that the element weights have uniform probability distribution with mean* μ *and variance* σ, *and* $\mathcal{J}_{max}$ *contains maximum sets of size* h_0, *and* $k \leq N^{-1}_{h_0\mu, h_0\sigma}(|\mathcal{J}_{max}|^{-1} - 1)$ *and also* $\text{WIR}_k^{(\ell)} = 1 + o(1)$ *as* $\ell \to \infty$. *Then*

$$\text{Prob}\{\exists I \in \mathcal{J}_{max} \textit{ with } W(I) \geq k\} \to 1 \quad \textit{as} \quad \ell \to \infty$$

where $N_{h_0\mu, h_0\sigma}$ *is the normal distribution function of mean* $h_0\mu$ *and variance* $(h_0\sigma)^2$. □

6. Intersections of RIS and Relation of RIS to Matroids

6.1 The Relation of RIS to Whitney's Matroids

Let us define for RIS M and $h \geq 0$

$$\lambda_M(h) = \text{Prob}\{M_0 \text{ is a matroid of rank } h \;/ M_0 \text{ is an instance of } M\} .$$

It is easy to establish a rough lower bound for $\lambda_M(h)$, given $M = (E, \mathcal{J}, p)$ is uniform. Let $\mathcal{J}_h = \{I / I \in \mathcal{J} \wedge |I| = h\}$.

PROPOSITION 6.1.

$$\lambda_M(h) \geq |\mathcal{J}_h| \cdot p^h (1-p)^{|E|-h} .$$

Proof. Note that for each $E_0 \in \mathcal{J}_h$, $M_0 = \{E_0, \{I \in I \subseteq E_0\}\}$ is an instance of M of probability $p^h(1-p)^{|E|-h}$ and M_0 is a matroid.

6.2 Intersections of RIS

Let M_1, M_2 be RIS with

$$M_1 = (E, \mathcal{J}_1, \{p_e^{(1)}\})$$

and

$$M_2 = (E, \mathcal{J}_2, \{p_e^{(2)}\}) .$$

We wish to consider independent sets in both $\mathcal{J}_1$ and $\mathcal{J}_2$.

Let $M_1 \cap M_2$ be the structure

$$M = (E, \mathcal{J}_1 \cap \mathcal{J}_2, \{p_e^{(1)} \; p_e^{(2)}\}) .$$

It is not difficult to show (by definition of proper RIS) that:

PROPOSITION 6.2. $M = M_1 \cap M_2$ *is a proper RIS if* M_1 *and* M_2 *are proper RIS.*

There is no known result relating the complexity of constructing maximum independent sets in random instances of $M_1 \cap M_2$ to the complexity of constructing a maximum independent set in random instances of $M_1 \cap M_2$. Although in practice we often have that if the extension-rotation algorithm succeeds with high probability on M_1 and M_2 separately then it succeeds with high probability on $M_1 \cap M_2$. (See [Reif, Spirakis, 1981]).

In contrast, *matroids are not closed under intersection.* The problem of constructing a maximal independent set in the intersection of k matroids has a polynomial time (in $|E|$) algorithm [Lawler, 1977] for $k = 2$, but it is known to be a NP-complete problem for any $k \geq 3$.

7. Conclusion

We have proposed here the RIS and the weighted RIS as a general combinatorial structure for formulating problems with random inputs. We found that our nonconstructive technique for testing the existence of maximum independent sets is broadly applied to a large range of problems with random inputs, which can be formulated as RIS.

A companion paper, [Reif, Spirakis, 1981], considers a randomized algorithm, (the *Extension-Rotation* algorithm) for efficiently constructing an independent set of size h_0 in an instance of a RIS. Given an

independent set I of size less than h_0, we attempt to *extend* I (by adding a new random element e to I) or else attempt to *rotate* I (by deleting an element e' of I and adding the new element e). The use of a rotation operation first appeared in Posa's [1976] existence proof for a Hamiltonian path in an undirected random graph of density O(log n/n). In [Reif, Spirakis, 1981] we provide a *general method of analysis* of the performance of the extension-rotation algorithm.

References

Cohen, J., J. Komlös, and T. Mueller, "The probability of an interval graph and why it matters," *Proc. Symposia in Pure Mathematics* 34, 1979.

Erdös, P. and A. Renyi, "On random graphs," *Publicationes Mathematicae* 6, 1959, pp. 290-297.

Erdös, P. and A. Renyi, "On the evolution of random graphs," *Publ. Math. Inst. Hung. Acad. Sci.* 5A, 1960, pp. 17-61.

Erdös, P. and J. Spencer, *Probabilistic Methods in Combinatorics*, Academic Press, New York, 1974.

Feller, W., *An Introduction to Probability Theory and Its Applications*, Vol. 1, Third Edition, John Wiley and Sons, New York, 1968.

Grimmet, G.S., and C.J. McDiarmid, "On coloring random graphs," *Math. Proc. Camb. Phil. Soc.* 77, 1975, pp. 313-324.

Glover, H. and J.P. Huneke, "Cubic irreducible graphs for the projective plane," *Discrete Mathematics* 13, 1975, pp. 341-355.

Karp, R.M., "The probabilistic analysis of some combinatorial search algorithms," *Algorithms and Complexity: New Directions and Recent Results*, J.F. Traub, ed., Academic Press, New York, 1976, pp. 1-19.

Korte, R. and D. Hausmann, "An analysis of the greedy heuristic for independence systems," *Annals of Discrete Mathematics* 2, 1978, pp. 65-74.

Kuratowski, K., "Sur le problème des courbes gauches en topologie," *Fund. Math.* 15, 1930, pp. 217-283.

Lawler, E.L., *Combinatorial Optimization: Networks and Matroids*, Holt, Rinehard and Winston, 1976.

Lekkerkerker, C.G. and J.C. Boland, "Representation of a finite graph by a set of intervals on the real line," *Fund. Math. Polska Akad.*

Lueker, G.S., "Maximization on graphs with edge weights chosen from a normal distribution," *Proc. Tenth Annual Symposium on Theory of Computing*, San Diego, California, 1978.

Matula, D.W., "On the complete subgraphs of a random graph," *Proc. 2nd Chapel Hill Conference on Combinatorial Math. and Its Applications*, University of North Carolina, Chapel Hill, May 1970, pp. 356-369.

Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, 1965.

Posa, L., "Hamiltonian circuits in random graphs," *Discrete Mathematics* 14, 1976, pp. 359-364.

Reif, J.H. and P.G. Spirakis, "Probabilistic analysis of random extension-rotation algorithms," to appear in 1981 appearing in preliminary form as last part of "Random Matroids," 12th Annual STOC, Los Angeles, 1980, pp. 385-397.

Tutte, W.T., *Introduction to the Theory of Matroids*, American Elsevier, New York, 1971.

Walkup, D.W., "On the expected value of a random assignment problem," draft, December, 1977.

Walkup, D.W., "Matchings in random regular bipartite graphs," draft, December, 1977.

Whitney, H., "On the abstract properties of linear dependence," *American J. Mathematics* 57, 1935, pp. 509-533.